

HLA/RTI 구현을 위한 분산 시스템 구조

Distributed System Architecture for Implementing the HLA/RTI

김 호경, 김 명선, 정 창성

e-mail : {hkim, kms}@snoopy.korea.ac.kr, csjeong@charlie.korea.ac.kr

고려대학교 전자공학과 병렬처리 연구실

TEL. (02) 927 - 7683

FAX. (02) 926 - 7620

요약

HLA(High Level Architecture)는 분산 시뮬레이션의 표준화를 위한 개념적인 분산 시스템 프레임워크이며, RTI(Run Time Infrastructure)는 HLA를 따르는 분산 시뮬레이션들을 위한 네트워크 하부구조이다. 분산 시뮬레이션은 분산 노드간에 효율적이고 인과 오류가 없는 통신 방법의 지원이 필수적이며, RTI는 시뮬레이션에 참여하는 분산 노드간 메시지 전송을 위하여 최선(best-effort) 전송 서비스와 신뢰(reliable) 전송 서비스를 규정하고 있다. RTI의 메시지 전송 서비스는 기본적으로 메시지 전송 과정에서 메시지간 인과 관계를 유지해야하는 다자간 통신 서비스이며, 본 논문에서는 인터넷상에서 인과 오류가 없는 다자간 통신을 위한 새로운 형식의 분산 시스템 구조를 제안하고, 성능 평가를 위한 실험 결과를 소개한다.

1. 서론

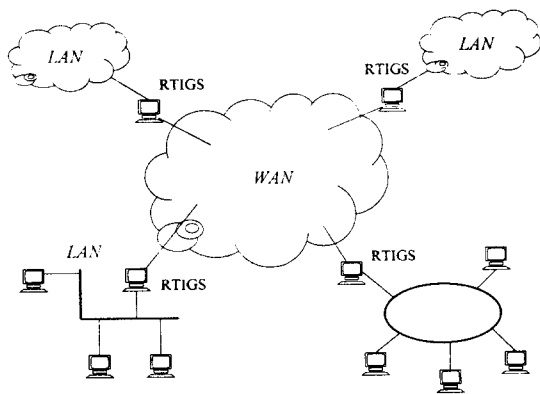
HLA는 미국 국방성에서 표준으로 제안한 분산 시뮬레이션 개발의 효율 향상을 위한 객체 기반 분산 시뮬레이션 프레임워크이며, 다음의 세 가지 구성요소로 이루어져 있다. : HLA 순용 규칙 (compliance Rules); 객체 모델 템플릿(Object Model Template); HLA 하부 구조 (Run-Time Infrastructure). HLA 순용 규칙[10]은 HLA 기반 시뮬레이션들이 지켜야할 논리적 규범이다. 또한, HLA는 페더레이션 객체 모델(Federation Object Model: FOM)과 시뮬레이션 객체 모델(Simulation Object Model: SOM)의 이원화된 객체 모델을 가지며, 객체 모델 템플릿[9]는 FOM 및 SOM의 서술을 위한 객체 모델 서술 도구이다. 마지막으로, HLA 하부구조(RTI)는 메시지 전송 및 객체 관리 등의 시뮬레이션 진행에 필요한 서비스를 제공하는 네트워크 하부구조이며, RTI에서 제공될 서비스의 종류 및 기능은 HLA 접속 사양 (Interface Specification)[11]에 기술되어 있다.

2 분산 시스템 구조

분산 시뮬레이션 진행 과정에서 각 노드들은 사건 정보를 메시지에 은닉하여 다른 노드들에게 전달하며, 메시지 처리 순서가 해당 메시지 내에 기술된 사건들의 인과 관계를 따르도록 함으로써 분산 환경에서 시뮬레이션을 수행할 때 생길 수 있는 논리적 오류를 방지할 수 있다. RTI에는 분산 노드간 메시지 전송을 위해 최선(best-effort) 전송 서비스와 신뢰(reliable) 전송 서비스가 정의되어 있으며, 각 노드에 도착된 메시지들이 수신 순서(receive order) 또는 타임스탬프 순서(timestamp order)에 따라 선택적으로 처리되도록 하고 있다.

위와 같이 RTI가 가지고있는 요구 수준을 만족시키기 위하여 본 논문에서 제안될 시스템 구조는 다음과 같은 구조적 특징을 갖도록 설계되었다:

- 본 시스템은 인터넷으로 연결된 서로 다른 네트워크상의 컴퓨터들을 단위 노드로 갖는 시뮬레이션 시스템 구축이 가능하도록 하기 위하여 네트워크 환경을 WAN과 LAN으로 구분하여 이중적으로 관리[2]하며, LAN에서는 UDP 브로드캐스트를 기본 통신 방식으로 하고 WAN은 UDP 유니캐스



[그림 1] 본 시스템의 네트워크 환경

트를 기본 통신 방식으로 사용한다. 또한, WAN과 LAN 사이를 오가는 메시지들에 대한 프로토콜 변환 및 메시지 필터링을 위하여 RTIGS(RTI Gateway Server)을 둔다. RTIGS는 이외에도 메시지 전송 순서를 보장해주고, 메시지 흐름 제어 및 네트워크 오류 검출 등의 기능을 담당한다. [그림 1]은 본 시스템의 네트워크 환경을 잘 예시하고 있다.

- 대규모 시뮬레이션을 지원하기 위해서는 시스템이 구조적으로 확장성(Scalability)을 갖추고 있어야 한다. 본 시스템은 RTIGS간 메시지 전송 트리 형태[4, 5]의 메시지 흐름을 갖도록 설계함으로써, 시뮬레이션 노드들의 수가 증가함에 따라 RTI에 의해 유발되는 네트워크 지연이 로그(log) 수준으로 최소화되도록 하였다.

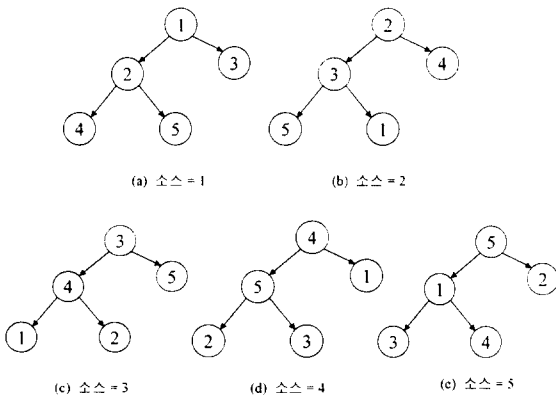
- 각 RTIGS들이 과도하게 WAN으로 메시지를 전송하는 것을 제한한다. 이는 WAN에서의 메시지 전송 지연이 상대적으로 LAN보다 크며, 중간 전송 과정에서 손실될 확률이 크기 때문이다. 또한, 메시지의 전송률을 제한하기 위하여 속도(rate)와 윈도우에 의한 흐름 제어 방법을 사용하였다.

3. RTIGS간의 메시지 전송

3.1 네트워크 구조와 전송 프로토콜

시뮬레이션 시스템에 참여하는 모든 RTIGS들은 시스템 내에서 고유한 ID를 가지며, 각 RTIGS에 할당된 ID를 이용하여 트리 형태의 메시지 전달 경로를 유지한다. [그림 2]는 5개의 RTIGS로 이루어진 시스템에서 각 RTIGS가 다른 RTIGS들에게 메시지를 전송할 때 사용하는 메시지 전송 경로를 잘 보여주고 있으며, 모든 RTIGS별로 고유한 메

시지 전송 경로를 가지고 있음을 볼 수 있다. 그러나, 각 RTIGS별 메시지 전송 경로는 별도의 자료 구조로 저장되는 것은 아니다. 대신에 각 RTIGS에 할당된 ID로부터 쉽게 찾아낼 수 있으며, [그림 2]에서 알 수 있듯이, 자신의 ID가 i 인 경우에 상위 노드의 ID는 각각 $(2*i)$ 와 $(2*i+1)$ 을 가지며, 상위 노드는 $(i/2)$ 를 갖는 완전이진(complete binary) 트리의 구조적 속성을 이용하였다.



[그림 2] 5 개의 RTIGS로 구성된 시스템의 네트워크 구조 예

각각의 RTIGS들은 시뮬레이션에 참여하는 RTIGS들에 대한 그룹 리스트를 가지고 있으며, 그룹 내에서 고유한 ID가 할당된다. 또한, 메시지 헤더에는 그 메시지를 최초로 전송한 RTIGS의 ID가 기록되어 있다. 그리고, 메시지를 수신한 RTIGS들은 메시지에 기록된 ID를 루트로 갖는 메시지 전송 경로 상에서 자신의 하위 RTIGS들을 계산해낸 다음, 하위 노드들에게 수신된 메시지를 다시 전송한다. 이 때, 메시지를 수신한 RTIGS는 메시지가 신뢰 전송 서비스를 요구하는 경우 부모 RTIGS에게 ACK 메시지를 보낸다. 이러한 전송 경로 트리는 각 RTIGS들이 메시지의 소스(source) ID와 자신의 ID만을 가지고 계산을 통해서 얻어내며, 별도로 트리를 물리적으로 저장하지

는 않는다.

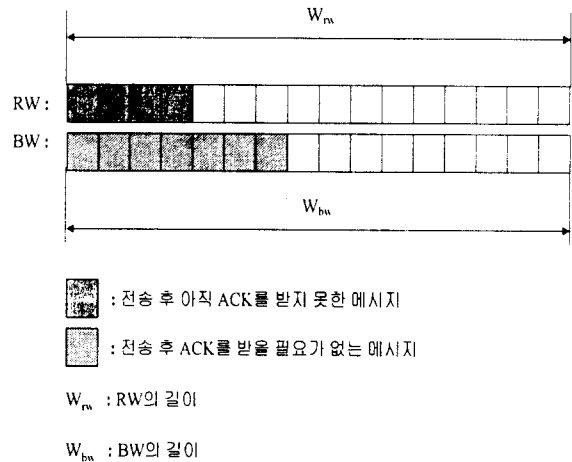
3.2 에러 제어

어떤 메시지가 신뢰성을 가지고 전송되어야 하는 경우, 전송 트리 내에서 하위 RTIGS들은 바로 위의 상위 RTIGS에게 수신된 메시지에 대한 ACK 메시지를 보냄으로써, 메시지 전송에 따른 신뢰성을 보장한다. 또한, 이 과정에서 제한된 시간에 상위 RTIGS에게 ACK 메시지가 도달하지 않은 경우에는 재전송을, 오류를 포함한 메시지를 수신한 하위 노드에서는 NAK(Negative ACK) 메시지를 상위 RTIGS에게 보내서 손실을 복구한다.

3.3 흐름 제어

본 시스템은 TMTP[5]와 RMTP[6]에서 사용된 속도(rate)와 윈도우(window)에 의한 흐름 제어 방법을 사용한다.

속도에 기반한 흐름 제어 방법은, RTIGS의 메시지 생성 속도가 정해진 임계치를 넘지 못하게 하는 방법이다. 이렇게 메시지 생성 속도에 최대값을 설정해서, RTIGS들에 의한 네트워크의 혼잡(Congestion)을 피하면서 일정 수준 이상의 QoS(Quality-of-Service)를 제공할 수 있다.



[그림 3] 메시지 전송 윈도우

속도 제어는 윈도우와 결합해서 이루어진다. 흐름 제어를 위해서 본 시스템에서는 두 개의 메시지 윈도우가 사용되며, 각각은 HLA 접속 사양에 규정된 신뢰전송 서비스와 최선전송 서비스를 위해 사용된다. 윈도우의 구조와 관련된 변수들은 아래 [그림 3]에 나와 있다.

위의 그림에서와 같이 신뢰전송 서비스가 필요한 메시지는 신뢰전송 윈도우(RW)에 들어가며, 최선전송 서비스가 필요한 메시지는 최선전송 윈도우(BW)에 들어가게 된다. 각 윈도우마다 T_{rw} , T_{bw} 의 시간 값이 설정되어 있어서, RW에 대해서는 T_{rw} 를 주기로 해서 윈도우내의 메시지를 검색해서 자식 서버들로부터 ACK 메시지를 받은 메시지를 제거하고 윈도우를 진행시킨다. BW에 대해서는, 여기에 있는 메시지들이 ACK 메시지를 필요로 하지 않기 때문에, T_{bw} 를 주기로 윈도우 내의 메시지를 제거하게 된다. 결국 메시지 전송률은 윈도우 크기와 각 윈도우에 설정된 시간 값에 의해 제한을 받게 된다.

4. 실험 결과

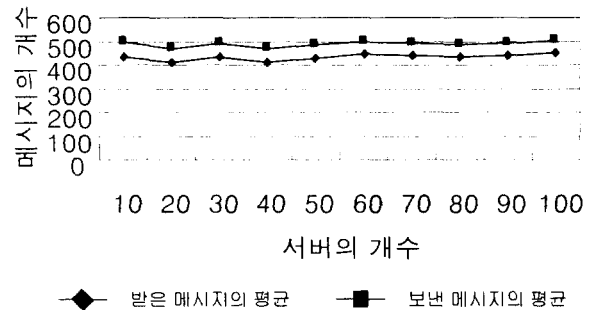
본 논문에서 제안된 시스템 및 프로토콜의 시험 환경은 다음과 같다.

- System : Linux, Pentium II Celeron 400MHz
- RAM : 128MB
- S/W tool : GNU gcc Compiler Version

2.7.2.3

실험은 한 개의 프로세스 내에 RTIGS역할을 하는 클래스와 채널(Channel)에 해당하는 클래스를 구현한 후, RTIGS의 개수를 변화시켜 가면서, 각각의 RTIGS 오브젝트들이 처리하는 메시지의 처

리량의 측정을 통해서 이루어졌다. 아래 [그림 4]에 나와 있는 것과 같이, 서버의 갯수가 증가하더라도 메시지의 처리량이 크게 증가하지 않는 것을 볼 수 있다. 아래 실험 결과를 통해서 제안한 시스템의 확장성(Scalability)을 확인할 수 있었다.



[그림 4] 메시지 처리량 변화

5. 결론 및 향후 연구

본 논문은 HLA의 네트워크 하부 구조인 RTI를 구현하는데 필요한 시스템 구조와 프로토콜을 제안하였다. 제안된 시스템은 인터넷 환경에서 메시지 전송 속도를 향상시키고 시스템의 확장성을 높이기 위해 WAN과 LAN의 이원화된 네트워크 구조를 사용하였으며, LAN과 WAN간의 연동은 RTIGS가 담당하도록 설계하였다. 또한, RTIGS간 메시지 전송 성능의 최적화를 위하여 완전이진 트리 형태의 메시지 전송 경로를 지원하며, 신뢰전송 서비스 및 최선전송 서비스가 효과적으로 지원될 수 있도록 설계되었다.

본 연구는 HLA에 대한 사양 검토를 시작으로, 현재는 제안된 시스템 구조 및 프로토콜에 대한 성능 분석이 진행중이며, 향후 프로토타입 구현을 위한 상세 설계가 진행될 계획이다.

[참고 문헌]

- [1] V. Jacobson, "Congestion Avoidance and Control," *Proceedings of ACM SIGCOMM'88*, Pages 314-328, August 1988.
- [2] Pullen, J. M and E. White, "Dual-Mode Multicast for DIS", 12th DIS Workshop on Standards for the Interoperability of Distributed Simulations, March 1995
- [3] H.W. Holbrook, S.K. Singhal and D.R. Cheriton, "Log-Based Receiver-Reliable Multicast for Distributed Interactive Simulation," *Proceedings of ACM SIGCOMM'95*, Pages 328-341, October 1995
- [4] S.Floyd, V.Jacobson, S. McCanne, C-G. Liu and L.Zhang, "A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing," *Proceedings of ACM SIGCOMM'95*, Pages 342-356, October 1995
- [5] R.Yavatkar, J. Griffioen, M. Sudan. "A Reliable Dissemination Protocol for Interactive Collaborative Applications", *Proceedings of ACM Multimedia*, Pages 333-344, 1995
- [6] J.C. Lin and S. Paul, "RMTP: A Reliable Multicast Transport Protocol," *Proceedings of IEEE INFOCOM '96*, Pages 1414-1424, March 1996
- [7] B.N. Levine, and J.J. Garcia-Luna-Aceves, "A Comparison of Known Classes of Reliable Multicast Transport Protocols," to appear in *Proceedings of International Conference on Network Protocols*, October 1996
- [8] B. Whetten, T.Montgomery, and S. Kaplan, "A High Performance Totally Ordered Multicast Protocol", available from research.ivv.nasa.gov ftp: /pub/doc/RMP/RMP_dagstuhl.ps
- [9] U.S. Defense Modeling and Simulation Office, "High Level Architecture Object Model Template Specification", Version 1.3, 5 February 1998
- [10] U.S. Defense Modeling and Simulation Office, "Department of Defense High Level Architecture Rules", Version 1.3, 5 February 1998
- [11] U.S. Defense Modeling and Simulation Office, "High Level Architecture Interface Specification Draft", Version 1.3, 25 April, 1998